DOCUMENT RESUME

ED 236 989 HE 016 760

AUTHOR Follman, John
TITLE Development of a University of South Florida Student Rating Scale.
PUB DATE 28 Oct 83
NOTE 6p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *College Faculty; *College Students; Course Evaluation; *Evaluation Criteria; Factor Analysis; Higher Education; Measurement Techniques; *Rating Scales; Student Attitudes; *Student Evaluation of Teacher Performance; Teacher Effectiveness
IDENTIFIERS *University of South Florida

ABSTRACT

The statistical results of administering a newly developed student rating scale of instructor effectiveness at the University of South Florida (USF) are analyzed, and the development and refinement of the scale are described. The scale was originally developed during the fall 1981 semester and was administered to 1,115 USF students. The scale contained five demographic items and 49 teacher/course rating scale items. Factor analysis reduced the 49 items to 10 interpreted factors. The scale was later refined to 34 instructor/course effectiveness items on a seven-category Likert rating scale. At the end of the spring semester, ratings on this scale were collected from 1,300 students of 14 instructors across four colleges. For this second statistical analysis, the student ratings were combined across instructors across colleges to provide a large sample and a university-wide perspective. Findings for each factor are discussed, and the following conclusions are made: the factor structures for the spring and fall semesters are consistent, and the ratings by USF students produced dimensions similar to those derived from students at other colleges. Four options for administrative and diagnostic uses of the scale are outlined. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED236989

Development of a University of South Florida Student Rating Scale

John Follman
University of South Florida

This is a report of the development and refinement of a student rating of instructor teaching effectiveness scale at the University of South Florida (USF), particularly the results of the statistical analyses. Initial development of this scale began during the fall semester of the 1981-82 academic year, when psychometrically sound, marker rating scale items, selected from the research literature, along with other items, were configured into a five category Likert Agree-Disagree rating scale and administered to 1115 USF students during the last two weeks of Semester I. The student rating scale then contained five demographic items and 49 teacher/course rating scale items. Early in Semester II that student rating scale data was factor analyzed which reduced the 49 items to 10 interpreted factors.

The rating scale was subsequently refined to 34 instructor/course effectiveness items on a seven category Likert rating scale, for additional data collection and analysis at the end of Semester II. Five demographic items were also included.

At the end of Semester II ratings on this scale were collected from ca. 1300 students of 14 USF instructors across four colleges, Arts and Letters, Engineering, Fine Arts, and especially Social and Behavioral Sciences. For this second statistical analysis, the student ratings were combined across instructors across colleges to provide a large sample and also a university-wide perspective.

AE016760

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John Follman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Principal components factor analysis and Kaiser Varimax rotation were conducted as in the Semester I analysis. The SPSS statistical package was used.

The principle components analysis produced six factors with eigenvalues in excess of one.

The first factor was a strong general factor with loadings in excess of .43 from 31 of the 34 rating scale items. This factor accounted for 45% of the variance. The remaining five factors respectively consisted of increasingly smaller numbers of items generally with increasingly smaller loadings, accounting for 7.2, and increasingly smaller amounts of variance, respectively.

The first group factor consisted of 10 items with loadings in excess of .60. Five items loaded greater than .77. These items, in order of strength, were: Instructor increased student interest; Instructor stimulated student intellectual curiosity; Instructor was excellent; Course was effective; and Amount learned was substantial. This factor is interpreted to represent an instructor/course positive student learning experience. The item loadings of items on this factor were strikingly similar to their loadings on its marker factor in the Semester I analysis.

The second factor consisted of three items: Instructor was prepared for class, loading .77; Instructor seemed to know the subject, loading .68; and instructor had everything going according to schedule, loading .62. This is an instructor structure-organization factor and is very similar to its predecessor.

The third factor had loadings in excess of .55 on six items. Major items were: Grading was fair and appropriate; Feedback on examinations was helpful; Instructor announced examinations in advance; and Instructor was fair and objective. This factor is interpreted to be examination and grading objectivity and fairness. This factor was partly consistent with its precursor in the initial analysis.

The fourth factor consisted of: Students frequently volunteered their own opinions, loading .76; Instructors complimented students on their work in front of others, loading .56; Students were able to argue with one another or the instructor, loading .55; and Instructor seemed to enjoy teaching this course, loading .53. This is a student participation factor, somewhat similar to its marker.

The fifth factor consisted of two items loading .84 and .83 respectively: Textbook and other readings were relevant; and Textbook and other readings helped integration. This textbook factor is almost identical to its predecessor in the first analysis.

The sixth factor was a tripleton of items: Instructor assigned very difficult readings, loading .80; Instructor assigned a great deal of reading, loading .71; and Instructor asked for more than students could get done, loading .70. This factor is clearly overload and is very similar to its analogue in the first analysis.

Overview across these second statistical analyses and also across the first statistical analysis suggests the following conclusions:

1) The factor structure derived in Semester II is generally consistent with that derived in Semester I.

2) The ratings by USF students of USF instructors produce empirically derived teaching effectiveness dimensions similar to dimensions empirically derived from other students at other institutions of higher learning.

3) Remaining are two developments. The first is the decision of the uses to be made, administrative, diagnostic, other. The second is the specification of the kinds of rating scales.

Following are options for the kinds of rating scales to be implemented:

a. One option, for adminstrative use, is a one-item overall instructor effectiveness scale (not recommended).

b. A second option, for administrative use, is items with large loadings, one large loading item each from each of the six empirically derived USF factors, one overall instructor effectiveness item, and one overall course effectiveness item. A draft of the items in this scale would include:

    1. instructor increased student interest
    2. examinations and grading were fair and objective
    3. reading assignments were demanding
    4. readings were relevant and helpful
    5. instructor was knowledgeable, prepared, and organized
    6. overall, instructor was effective, and
    7. overall, course was effective

c. A third option, for faculty self-improvement use, not for administrative use, is a scale of 20-30 items. Available upon request is a table of 27 items with their loadings from the statistical analyses from both semesters. These items all load high enough and consistently enough to empirically justify their use. Different faculty, departments, colleges can then select items they deem appropriate, cafeteria style.

d. A fourth option, for a research dependent variable, is a psychometrically super strong scale of two or more items per concept. This scale will be evaluative rather than descriptive.

Finally, I recommend that all scales for administrative or diagnostic use, include, in addition, open-end items appropriate to the particular department, college, university, etc.

10/28/83
CPT:EA1-3